Information

Sleep Deprivation and Cognitive Testing in Internal Medicine House Staff

JOHN ROBBINS, MD, *and* FRED GOTTLIEB, MD, *Sacramento, California*

There is increased concern about the effects of sleep deprivation on physician performance. We administered four standard tests of cognitive function to 23 university hospital house staff. Each physician served as his or her own control, and the tests were administered at rest, after a night on call, and after a night of sleep for recovery. The study was designed so that normal learning would minimize any deterioration in the post-on-call test performance. Statistically significant deterioration occurred in 3 of the 4 tests after a night on call. Even physicians acclimated to sleep deprivation on a regular, every-third-or-fourth-night basis showed functional impairment. The results have implications for patient care under conditions where house staff are stressed by sleep deprivation and prolonged fatigue.

(Robbins J, Gottlieb F: Sleep deprivation and cognitive testing in internal medicine house staff. West J Med 1990 Jan; 152:82-86)

Physicians in residency training programs are required to work for prolonged periods under conditions of sleep deprivation. It is at times of most severe sleep deprivation, that is, "on-call" nights, that residents are most likely to be responsible for the care of acutely ill patients. The therapeutic interventions required at such periods necessitate a high level of psychomotor performance, cognitive function, and emotional equilibrium. The greatest demand on a physician's skills often comes when performance is most likely impaired by sleep loss.

Despite the large amount of research in the field of sleep deprivation, the data regarding physician performance under sleep-deprived conditions are sparse. Existent information implies potentially adverse effects on mood and behavior.[1-8] Although there are notable exceptions,[4,6,9,10] few studies specifically assess the psychomotor and cognitive performance of house staff under conditions of acute sleep deprivation. The classic medical study was published almost 20 years ago when Friedman and co-workers showed that sleep-deprived house staff made more errors when interpreting electrocardiograms.[9] That study addressed one important cognitive aspect of medical practice. There may be more global deterioration of post-on-call house-staff performance.

Deaconson and associates were unable to show the effect of sleep deprivation on surgical residents' performance on psychometric tests.[10] That study used residents on an every-other-night call schedule as their own controls. As pointed out in numerous letters to the editor that followed the original publication, residents in the "control state" may also have been considerably impaired from chronic sleep deprivation.[11]

To better understand the effects of acute sleep deprivation on house-staff performance, we compared the performance of house staff on standardized cognitive tests when rested, after being on call—that is, acute sleep deprivation—and a third time after conditions of normal sleep. The standard on-call schedule at the institution is every fourth night. The control, or rested, evaluations all took place when the resident had been able to get a full night's sleep at home before taking the test. The study was designed in such a way that if previous sleep deprivation had an effect on the results, it would mitigate against the finding of a statistically significant outcome. The session format allowed us to measure baseline functioning and functioning after sleep deprivation and to recheck the baseline, taking into account the normal learning curve.

Methods

Standard tests routinely used in cognitive assessment and rehabilitation were selected to evaluate various cognitive functions. They were administered on a microcomputer terminal with the investigator present at every testing period. The four tests were as follows*:

- *Driver,* in which the subject must maneuver a "car" on a moving track of set width and speed, using a joystick.[12] Scoring is based on percentage accuracy and increased by keeping the car at the bottom of the screen—that is, with less lead time before road hazards appear, which minimizes reaction time. The final score is, therefore, expressed as percentage accuracy and total score. Attention span, eye-hand coordination, and reaction time are assessed.
- *Rapid Number Comparison,* in which four numbers of five digits each are flashed on the screen for approximately 0.5 seconds.[13] Two numbers are identical, and that number

*Bonnie Napier, OT, assisted with the computer programs and testing.

From the Department of Internal Medicine, Division of General Medicine, University of California, Davis, Medical Center, Sacramento, California.
Dr Gottlieb is currently with the Northwest Community Health Center, Salt Lake City.
Reprint requests to John Robbins, MD, Division of General Medicine, 3120 Primary Care Center, 2221 Stockton Blvd, Sacramento, CA 95817.

must be accurately recognized and logged into the computer. Visual reaction time and accuracy, rapid numeric discrimination, short-term memory, and attention span are assessed.

- *Concentration,* in which the subject is presented with a field of characters arranged in a random pattern.[14] A character is selected sequentially by the computer as the target. After the target is located, the subject maneuvers a cursor to cover the character. Scoring is based on the speed of locating the selected character and moving the cursor minus incorrect responses. If a target is not located in a specified time, the subject must start over. Spatial perception, performance under prolonged repetitive conditions, the ability to discern relevant from random information, eye-hand coordination, and the ability to perform under time limitations are assessed. Assistance was given to complete the test to residents who were unable to do this test in the allotted time under post-on-call conditions. Final scoring was affected only in that it allowed test completion for data interpretation.
- *Smart Shopper Marathon,* referred to hereafter as *Math Problems,* is a test in which a subject is presented five sets of simple, high-school level mathematic problems under time limitations.[15] Scoring is based on the number of correct answers in each problem set. There is a time penalty for incorrect answers. Each set requires a minute to complete, and simple mathematic abilities under time limitations are assessed.

The testing period required for these four tests is about 30 to 40 minutes. Each subject who agreed to participate was given a baseline test under conditions of a normal night of sleep for that person. A second test was given under conditions of sleep deprivation. At this second testing period, the volunteering house officer would page the investigator at his or her own convenience before leaving the hospital on completing call night duties (approximately 35 hours from the end of their last normal sleep period). A third and final testing period was given at the convenience of each subject, again under non-sleep-deprived conditions, to determine if change and performance trends under conditions of sleep deprivation would persist after the subjects were rested. Therefore, a change noted in overall performance under conditions of sleep deprivation should revert to baseline if sleep deprivation is the only variable.

TABLE 1.—*Cognitive Testing of Sleep Deprivation—Correlation of Scores* (r) *in Repetitive Testing (Reproducibility)*

	Before On-Call Night	*After On-Call Night*	*After Recovery*
% Accuracy	0.9277	0.9298	0.8860
Total score, *r*	0.8641	0.9091	0.9495

To assess the reproducibility of results and the validity of these tests, the test with the greatest potential range of scores, *Driver,* was administered a second time at each testing session to 17 of the 23 subjects who agreed to donate the additional time required.

Before doing the test under conditions of sleep deprivation, each subject estimated the amount of sleep obtained on call and the total number of hours since getting a full night's sleep. This is obviously subjective but was intended to make sure none of the subjects had spent two successive nights in a sleep-deprived state, but that all had some degree of sleep deprivation. Baseline assessments were carried out at various times during the normal working day.

In analyzing data, each subject served as his or her own control. No incentive was offered for participation or performance in the study.

Subjects

Members of the house staff in the internal medicine department at the University of California, Davis, Medical Center (Sacramento, California) were asked to volunteer in a study to determine the effects, if any, that sleep deprivation typical of a night on call might have on cognitive performance. In all, 31 of the house staff completed the first battery of tests under conditions of normal sleep. Eight house staff completed the baseline study but were unable to complete the remainder of the study. Of this group, five reported being too tired after being on call and three were not on rotations requiring being on call during the ten-week period of this study. Thus, 23 house staff completed the study. The male-to-female ratio was 16:7, and the distribution of year in training—9 in their first year after graduating medical school, 7 in their second year, 6 in their third year, and 1 in

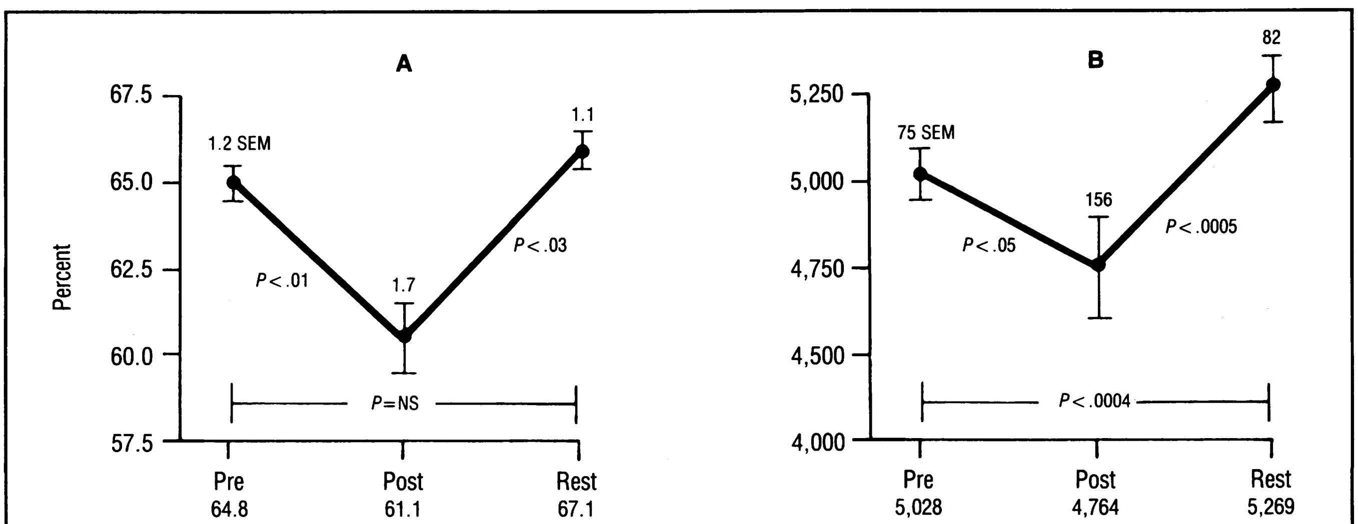

Figure 1.—The graphs show performance results on *Driver* tests before (Pre) and after (Post) being on call and when rested (Rest), with **A** showing the percent accuracy and **B** the total scores indicating speed of response. NS = not significant, SEM = standard error of the mean

the fourth year—represented a valid cross section of the house staff in the Department of Medicine at this university teaching hospital.

Results

Tests were administered under sleep-deprived conditions, averaging 2.1 hours of sleep ± 1.0 hours standard error of the mean (SEM). The total hours awake since the start of on call were 33 hours ± 2.6 SEM. All tests assessing performance under conditions of sleep deprivation were given in the afternoon following on call from 1 to 7 PM.

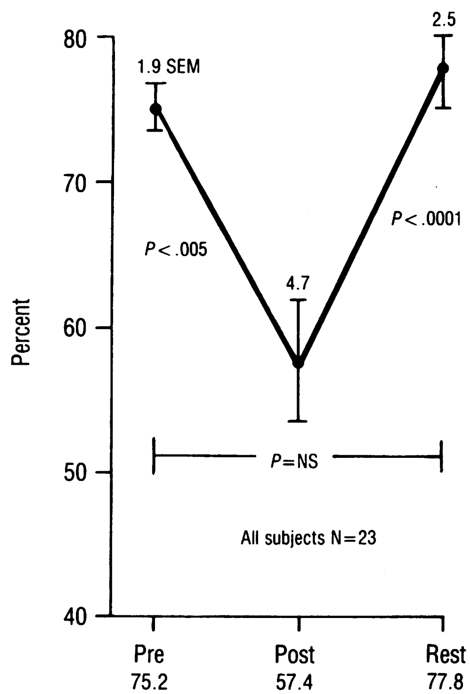

Figure 2.—The graph shows performance results (%) on the *Rapid Number Comparison* test when given before (Pre) and after (Post) being on call and when rested (Rest). NS = not significant, SEM = standard error of the mean

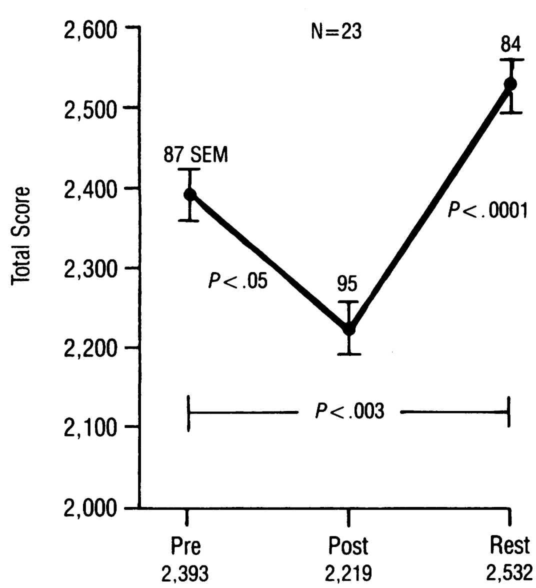

Figure 3.—The graph shows performance results on the *Concentration* test when given before (Pre) and after (Post) being on call and when rested (Rest). SEM = standard error of the mean

Reproducibility

Pearson's correlational analysis was performed on the scores of the 17 subjects who completed the *Driver* test twice in each testing period. As can be seen in Table 1, there is a high correlation between repetitive performances on the same test administered immediately after an initial attempt, with *r* values ranging from .86 to .95.

Results of Each Test

Paired *t* tests were used to analyze the results.

The percent accuracy on *Driver* testing decreased from that before an on-call night to that after being on call, with a significant deterioration in performance ($P < .01$). The improvement from testing after an on-call night to the recovery at rest was significant ($P < .03$). There was no change from the initial test to recovery on the percentage scores. On the raw score data, which included accuracy plus speed of response, the deterioration from before to after on-call testing was significant ($P < .05$). The improvement from the post-on-call test results to the results after recovery was significant ($P < .0005$). As the graph shows (Figure 1), there is a significant improvement in the third performance results under rested conditions from those of the first test under rested conditions ($P < .004$).

Rapid Number Comparison testing showed a significant deterioration under post-on-call conditions ($P < .005$). Improvement after sleeping was also significant ($P < .0001$). There was no significant change from the first test under rested conditions to the final control test under rested conditions (Figure 2).

Concentration testing showed a deterioration in the results under post-on-call conditions significant to the *P* less than .05 level. There was significant improvement after rest compared with the post-on-call conditions ($P < .0001$). There was significant improvement from the rested recovery test results compared with those of the baseline test ($P < .003$) (Figure 3).

There was an incremental improvement with each at-

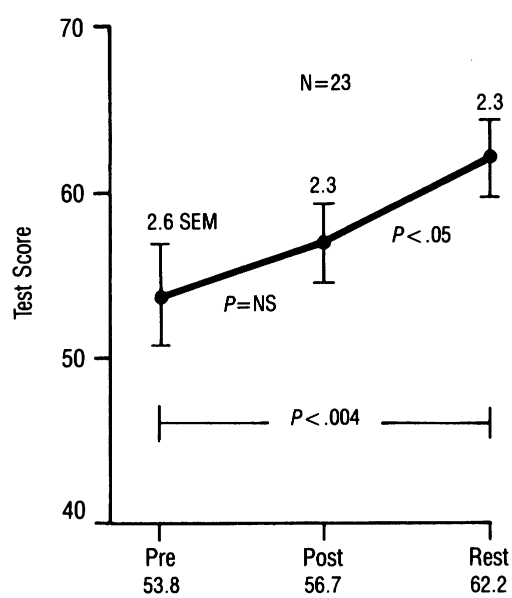

Figure 4.—The graph shows performance results on the *Math Problems* test when given before (Pre) and after (Post) being on call and when rested (Rest). NS = not significant, SEM = standard error of the mean

tempt in *Math Problems,* but this trend was not statistically significant. There was improvement in the results from the third test under rested conditions compared with after being on call, which was significant ($P < .05$). The initial test results compared with those of the third test under rested conditions did show significant improvement ($P < .004$) (Figure 4).

An attempt was made to identify specific factors that would predict deteriorating performance when sleep deprived. No significant correlation could be identified between change in performance and hours of sleep obtained on call, total hours awake, age, sex, or year in training and hospital rotation of subjects.

Discussion

Several factors included in the design of this study minimize its likelihood of demonstrating a statistically significant deterioration on these cognitive tests. First, the possible effect of practice or learning on these tests decreases the difficulty of the post-on-call test. The effect of learning can be present both in the familiarity with the test and with increasing dexterity in manipulating the microcomputer keyboard. Tests did not vary, the mathematic problems were identical at each test, and residents learned various shortcuts available in maneuvering the cursor. It is likely that learning played some role in the scoring on these tests. The degree to which this may have modified the results cannot be estimated, but it would serve to minimize the effect of sleep deprivation on performance because each successive test period should show some improvement purely from learning. Certainly, learning may best explain the improvement seen under the rested conditions as compared with the results on the preliminary rested tests.

Second, there was an inherent bias in the subjects who participated and completed the test. Of the initial 31 subjects who participated in the baseline study, 5 did not complete the study because they said they were too tired. Therefore, the results are modified by including only those residents who were moderately fatigued but still able and willing to participate and complete the study. Those most exhausted were self-excluded.

Third, assistance was given on one test to those too fatigued to complete the test on their own. In summary, all possible controllable factors in the design of the study were set to work against the hypothesis that cognitive performance would deteriorate after sleep deprivation.

Various studies have shown a deterioration in performance when sleep is limited to less than three[6] to five[16,17] hours a night. Our experience from this institution and at other teaching hospitals is that house staff rarely get more than three hours of sleep per night when on call.

Previous literature has also shown that the more complex the task, the greater the deterioration likely to be evident under conditions of sleep deprivation.[4] Similarly, the greater the time required to complete the test, the more likely deterioration would be evident. A commonly stated period of time for testing before which deterioration is less likely to be evident is about three minutes.[4] This is supported by the results of this study, which showed deterioration in the three more complex tasks presented by the *Driver, Concentration,* and *Rapid Number Comparison* tests. The simple *Math Problems* required testing periods of less than a minute each. This test showed no statistically significant deterioration under post-on-call conditions.

Research has also shown that visual vigilance is one of the more highly sensitive areas to show deterioration in performance following sleep deprivation. This was confirmed in our tests showing greater deterioration in rapid number recognition.[4,7,9] Other studies have shown that impairment in task performance following sleep deprivation persists the day following sleep.[18] Residents who routinely go without substantial sleep every third or fourth night still showed significant deterioration.

A question could be raised as to whether the residents had a "secondary gain" in demonstrating deterioration after being on call and therefore might not give an optimal performance on the post-on-call test. As with pulmonary function tests, it is extremely difficult to reproduce suboptimal results in a consistent manner on these tests. The high degree of correlation on repeat testing in all three sessions argues against less-than-maximal performance. Furthermore, residents are generally highly motivated to provide the best services for the patients entrusted to their care. To suggest less than optimal effort would be consistent with emotional imbalance reported elsewhere.[3,7,8] That this imbalance would have its effects on the entire subject population would add further evidence to the detrimental effects of sleep deprivation on house-staff affect during on-call nights.

Internal medicine residents even on an every-fourth-night schedule may suffer from chronic sleep deprivation. If this is the case, this study would have a decreased likelihood of showing significant results. Any decrement in baseline, control function should have decreased the divergence between the sleep-deprived and control periods. Had the subjects been completely well rested, such as from no sleep deprivation for months, we could assume that results would have shown more notable differences in performance.

No actual norms exist for the tests used. They are not intended to test actual function but rather an improvement or deterioration in individual function. Previous studies have looked at physicians' ability to perform specific medical tasks.[9] The strength of this battery of tests is that it tests more global functioning in sleep-deprived medical house officers. The more complex tasks showed greater deterioration than the simple tasks. Hand-eye coordination and the ability to process complex data were more affected than simple tasks.

An argument can be made that this study did not directly assess medical skills such as placing arterial lines, interpreting blood gas measurements, or making rapid diagnoses. The various tests assessed cognitive function ranging from reaction time and eye-hand coordination to attention span and rapid numeric discrimination, which has direct implications on those skills required of practicing physicians. We chose to use standardized cognitive tests to elicit and detect a change in psychomotor and cognitive functions. That the duties and responsibilities of direct patient care require far more complex cognitive functions than these tests can hope to assess cannot be denied. The deterioration evident in these tests may therefore be moderate relative to the actual effects sleep-deprived conditions may have on cognitive tasks required for direct patient care.

This study showed a deterioration in performance in cognitive testing after periods of short-term sleep deprivation—that is, on-call nights—that are standard in many training programs. If further studies confirm this relationship between sleep deprivation and a decline in cognitive performance, this may have medicolegal implications for patient care under conditions where house staff are deprived of normal sleep. Training programs are going to have to change. This study confirms what common sense assumes: house staff perform considerably less well after sleep deprivation. The occupational hazard of prolonged sleep deprivation may also have further implications for residents in house-staff training programs who drive home after an on-call night. Given the results of this study, we are concerned about ensuring optimal patient care and resident safety after a sleepless, stressful night on call.

REFERENCES

1. Asken MJ, Raham DC: Resident performance and sleep deprivation: A review. J Med Educ 1983; 58:382-388
2. Ford CV, Wentz DK: The internship year: A study of sleep, mood states, and psychophysiologic parameters. South Med J 1984; 77:1435-1442
3. Friedman RC, Kornfeld DS, Bigger TJ: Psychological problems associated with sleep deprivation in interns. J Med Educ 1973; 48:436-441
4. Hawkins MR, Vichick DA, Silsby HD, et al: Sleep and nutritional deprivation and performance of house officers. J Med Educ 1985; 60:530-535
5. McCue JD: The distress of internship—Causes and prevention. N Engl J Med 1985; 312:449-452
6. Poulton FC, Hunt GM, Carpenter A, et al: The performance of junior hospital doctors following reduced sleep and long hours of work. Ergonomics 1978; 21:279-295
7. Sapira JD: Dysphoria and impaired mentation in young physicians. South Med J 1977; 11:1305-1307
8. Small GW: House officer stress syndrome. Psychosomatics 1981; 22:860-869
9. Friedman RC, Bigger JT, Kornfeld DS: The intern and sleep loss. N Engl J Med 1971; 285:201-203
10. Deaconson T, O'Hair DP, Levy MF, et al: Sleep deprivation and resident performance. JAMA 1988; 260:1721-1727
11. Sleep deprivation and performance of residents (Letters). JAMA 1989; 261:859-863
12. Bracy O: Driver (Software Program). Cognitive Rehab 1983; 1:23-24
13. Engum ES: Rapid number comparison (Software Program). Cognitive Rehab 1984; 2:15-18
14. Cockcroft P: Apple/Concentrate (Software Program). Fam Comput 1984; 13:110
15. Koetke W: Smart Shopper Marathon, *In* Survival Math (Software Program). New York, Sunburst, 1985
16. Hartley LR: A comparison of continuous and distributed reduced sleep schedules. Q J Exp Psychol 1974; 26:8-14
17. Sassin JF: Neurological findings following short-term sleep deprivation. Arch Neurol 1970; 22:54-56
18. Wilkinson RT: Aftereffect of sleep deprivation. J Exp Psychol 1963; 66:439-442